In Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2008), Vancouver, Canada, August 20-24, 2008 (In Press).

Comparison of Filtering and Classification Techniques of Electroencephalography for Brain-Computer Interface

Mark Renfrew[*], Roger Cheng[†], Janis J. Daly[‡], M. Cenk Cavusoglu [§]

Abstract

In this paper several methods are investigated for feature extraction and classification of mu features from electroencephalographic (EEG) readings of subjects engaged in motor tasks. EEG features are extracted by autoregressive (AR) filtering, mu-matched filtering, and wavelet decomposition (WD) methods, and the resulting features are classified by a linear classifier whose weights are set by an expert using a-priori knowledge, as well as support vector machines (SVM) using various kernels. The classification accuracies are compared to each other. SVMs are shown to offer a potential improvement over the simple linear classifier, and wavelets and mu-matched filtering are shown to offer potential improvement over AR filtering.

1. Introduction

A brain-computer interface (BCI) is a system for direct communication between a human or animal and a computer. The computer uses a subject's brain signals as input, processes them by finding and classifying features present in the signal, and performs some action based on its classification, such as moving a cursor [1]. Improvement in data collection equipment and computer processing speed has led to an increase in BCI research in recent years. Much of this research has been aimed at restoring autonomy to patients who have severely reduced motor control, such as those with Amyotrophic Lateral Sclerosis. We are interested in the possibility that BCI may be effective as a therapy tool for the motor rehabilitation of stroke patients, by facilitating targeted learning in the areas of the brain damaged by stroke.

Brain signals are often collected by electroencephalography (EEG), which is done by placing electrodes on the subject's scalp and reading the voltages produced by cortical activity. EEG has the advantages of being noninvasive and relatively easy to perform, but is more susceptible to noise from muscular contractions and has a lower spatial resolution than invasive methods such as electrocorticography (ECoG) [2]. The data used in this study was collected using scalp EEG.

Since we are concerned with BCI as a means of stroke motor rehabilitation, we use the cortical mu rhythm as the feature of interest. This is a signal feature that is present in the EEG of most healthy adults, especially over motor areas of the brain [3]. The mu rhythm is an arch-shaped oscillation that is strongest in the 8 – 13 Hz range (the alpha component), but is also present from 13 – 30 Hz (beta) and $>$ 30 Hz (gamma). Mu rhythm is attenuated by motor activity, a phenomenon known as event-related desynchronization (ERD). Additionally, most people can be trained to have a great deal of control over their mu rhythms [4].

Detection of the mu rhythm is complicated by the fact that its frequency overlaps with the alpha rhythm, which occurs from 8 – 12 Hz and is associated with visual stimulation, and the beta rhythm, which occurs above 12 Hz and is associated with waking consciousness. It is difficult to separate the mu rhythm from the alpha and beta rhythms based solely on spectral information, but the spiky shape of the mu rhythm may be exploited in order to detect its presence. In this paper, we utilize two techniques in an attempt to exploit this: a matched filter for mu waves, and wavelet decomposition (WD), which is known to be more effective at detecting sharp, spiky signal features than traditional spectral techniques [4, 5]. We compare the classifica-

[*]Mark Renfrew (mark.renfrew@case.edu) and M. Cenk Cavusoglu (cavusoglu@case.edu) are with the Department of Electrical Engineering and Computer Science, Case Western Reserve University, Cleveland, OH

[†]Roger Cheng (rcc13@case.edu)is with the Cognitive and Motor Learning Laboratory, Louis Stokes Cleveland Dept. of Veterans Affairs Medical Center, Cleveland, OH.

[‡]Janis J. Daly (jjd17@case.edu) is Director, Cognitive and Motor Learning Laboratory, Louis Stokes Cleveland Dept. of Veterans Affairs Medical Center, Cleveland, OH, and with the Department of Neurology, Case Western Reserve University School of Medicine, Cleveland, OH.

[§]This work was supported by the Department of Veterans Affairs, Office of Rehabilitation Research and Development grants B5080S & B3709R

tion accuracies obtained with these techniques to those obtained using autoregressive (AR) filtering.

After extracting features from EEG, a BCI system must decide what they mean. That is, it must assign a classification to each sample based on the features calculated for that sample. We investigate the performance of two classification techniques: a linear combination of the EEG features with a constant weight vector that is determined by an expert, and support vector machines (SVM) with several different kernel functions.

2. Methods

2.1. Task

The EEG data analyzed in this study was collected during BCI therapy sessions. In these sessions, all subjects were seated in front of a computer screen with their right hand gripping the manipulandum, or end-effector, of an IMT Inmotion2 shoulder-elbow robot. EEG data was obtained using 58-channel ECI ElectroCap EEG cap and Compumedics Neuroscan software and amplifiers. The electrode locations on the cap conformed to the International 10-20 standard and all electrodes were referenced to ground electrodes on the ears. All electrode-scalp impedances were reduced to under 5kΩ by use of electrically conductive gel and the impedance-measurement facilities provided by the Neuroscan Acquire software. EEG data was sampled and digitized by Neuroscan Acquire, with a gain of 500, a sampling rate of 250 Hz, and bandpass filtered from 0.1 – 40 Hz.

This data was sent to the input module of BCI2000, an open-source modular BCI implementation [6]. Subjects performed a modified version of the BCI2000 D2Box task, in which two targets (up or down) were presented onscreen in random order. Subjects were asked to relax when the bottom target was presented, and to alternately perform a real or imagined reaching motion when the top target was presented. The trials were of two types: screening, which were used to calibrate the classifier settings and in which no cursor was moved onscreen, and training, in which the BCI2000 signal processing module performed filtering and classification of the EEG data in order to move a cursor onscreen with which the subject attempted to hit the target. Each trial consisted of a series of target presentations, terminating after 180 seconds had elapsed or the subject hit 10 targets with the cursor.

The BCI data from two healthy right-arm dominant subjects, ages 25 and 20, were collected for analysis. Only data from imaginary trials were analyzed in this study, in order to avoid any possible EMG contamination induced by the reaching movement. Data collection was conducted according to the Declaration of Helsinki and oversight was provided by the Internal Review Board of the Louis Stokes Cleveland Veterans Affairs Medical Center. The data were processed as follows.

2.2. EEG Processing

The EEG data were preprocessed by using a common-average reference (CAR) spatial filter to reduce noise [7]. This is performed by applying the formula

$$V_i^{CAR} = V_i^{raw} - \frac{1}{58}\sum_{j=0}^{58} V_j^{raw} \tag{1}$$

where V_i^{raw} is the potential between the ith channel and the reference channel, and V_i^{CAR} is the spatially filtered value for the ith channel.

The three methods of temporal filtering mentioned above were then performed on V^{CAR}: AR filtering, which is the method that is used by BCI2000 used in the online trials, mu-matched filtering, and wavelet decomposition.

AR filtering estimates a signal's spectral density with the equation

$$\hat{P}(e^{jw}) = \frac{1}{\left|1 - \sum_{k=1}^{p} a_p(k) e^{-jkw}\right|^2} \tag{2}$$

where $a_p(k)$ are time-varying filter coefficients which are estimated by an AR process, and p is the AR model order [8]. The result is that P is a series of numbers that give the strength of the signal in various frequencies.

The spatially filtered EEG signals were processed with a 12th order AR model, which has been shown to be the optimal order to extract information from the EEG alpha band for BCI [8]. 10 spectral estimates were obtained, each representing the power of a 3 Hz slice of the spectrum from 0 – 30 Hz.

Mu-matched filtering was done by comparing the signal with an approximation of the canonical mu rhythm, which is a sharp rectified sinusoid defined by

$$s_n(n) = h\left|sin\left(\frac{n\pi f_F}{f_S} + \frac{m\pi}{K}\right)\right|,\ m = 0, 1, ..., K$$

$$h_l s(x) = \frac{1}{1+e^{-Ax+B}} \tag{3}$$

where n is the sample number, f_S is the sampling frequency, f_F is the frequency of the template, and A, B, and K are experimentally determined parameters.

Wavelet decomposition of a signal X is done by first choosing a wavelet function ψ, which has four filters associated with it: a high-pass decomposition filter G, a low-pass decomposition filter H, a high-pass reconstruction filter G', and a low-pass reconstruction filter H'. Then the convolution between X and the filters G and H is computed, giving two sets of coefficients. Both these sets of coefficients are decimated by a factor

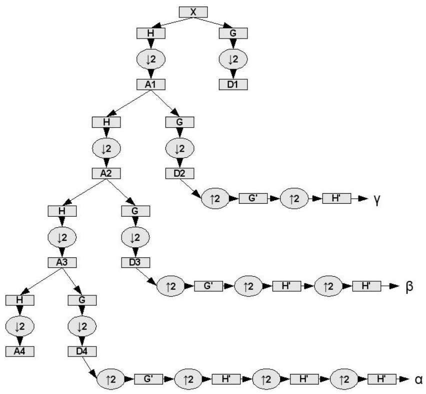

Figure 1. ***The WD scheme used to extract the alpha (8 – 16 Hz), beta (16 – 31 Hz) and gamma (31 – 62 Hz) components of EEG.***

of two to remove redundant information. This produces signals D, which carries the high-frequency information of X, and A, which carries the low-frequency information. The process may be repeated recursively on D or A to extract desired frequencies. X can be reconstructed exactly by upsampling D and A (i.e., inserting a zero after every sample), and convolving with the reconstruction filters G' and H' and then summing [9].

The EEG data was sampled at 250 Hz and so by Nyquist's rule carries frequencies from 0 – 125 Hz. Alpha, beta, and gamma components of EEG can then be extracted using a 4-level decomposition and reconstruction scheme, as shown in Fig. 1. Four wavelets of two different families were tested in this study: Biorthogonal 4/4 wavelets and Daubechies 2nd, 8th, and 25th order wavelets.

2.3. Classification

The first classification technique, which is the one used by BCI2000 in the online task, classifies samples by applying the formula

$$(\mathbf{x}) = sign(\mathbf{w}^T \mathbf{x} + b) \tag{4}$$

where $\mathbf{x}$ is the vector of EEG features, $\mathbf{w}$ is constant weight vector chosen by an expert, b is a bias term, and f is the classification value.

We then use a support vector machine (SVM), which is a learning algorithm that maximally separates the samples of distinct classes by solving the equation

$$(\mathbf{x}) = sign(\mathbf{w}^T \phi(\mathbf{x}) + b) \tag{5}$$

where ϕ is a function that maps $\mathbf{x}$ into some possibly high-dimensional space. This is known as the *kernel trick*, and can be exploited to use the SVM as a nonlinear classifier [10].

We use LIBSVM, an open-source SVM library, to train SVMs on the screening trials using linear,

Class. Method	mean	std.dev	p-value
db25 / SVM poly.	63.11	9.44	< 0.01
db8 / SVM sig.	62.57	7.50	< 0.01
db25 / SVM lin.	62.46	7.65	< 0.01
db8 / SVM RBF	62.38	8.14	< 0.01
match / SVM lin.	61.85	9.53	< 0.01
db8 / SVM poly.	61.40	7.90	< 0.01
match / SVM poly.	60.70	11.16	< 0.01
AR / SVM RBF	60.70	10.41	< 0.01
db25 / SVM sig.	60.68	8.41	< 0.01
db2 / SVM RBF	60.47	7.36	< 0.01
db2 / SVM sig.	60.45	8.38	< 0.01
db8 / SVM lin.	59.33	6.40	< 0.01
AR / SVM poly.	59.27	11.78	< 0.01
AR / SVM lin.	59.17	8.16	< 0.01
db25 / SVM RBF	59.16	11.12	< 0.01
bior44 / SVM sig.	59.09	6.40	< 0.01
bior44 / simple lin.	58.86	7.53	< 0.01
bior44 / SVM poly.	58.85	11.56	< 0.01
match / SVM RBF	58.33	10.30	< 0.01
match / SVM sig.	58.04	9.95	< 0.01
match / simple lin.	57.54	6.55	< 0.01
db2 / SVM lin.	57.25	8.19	< 0.01
db2 / simple lin.	55.70	11.66	0.23
db8 / simple lin.	55.32	11.17	0.30
bior44 / SVM RBF	55.15	12.92	0.40
db2 / SVM poly.	54.56	11.95	0.54
bior44 / SVM lin.	54.25	11.46	0.63
AR / SVM sig.	54.23	11.87	0.64
AR / simple lin.	**53.28**	**6.50**	-
db25 / simple lin.	52.22	12.31	0.61

Table 1. ***Sorted mean classification accuracies and standard deviations (in percent) of all classifier / signal processing method pairs tested, and probability of statistical similarity to the baseline method of AR / simple linear classification(in bold).***

polynomial, radial basis function, and sigmoid kernels [11, 12]. Training sets consisted of data collected during three three-minute screening sessions. To keep training time to a reasonable level, training sets were reduced to 2000 randomly selected samples, with an equal number being from each class. Testing sets consisted of data collected in 22 therapy sessions. Each testing session is considered a trial. Classification accuracies for each method are compared against each other, and the statistical significance of the differences were calculated using a one-way ANOVA test in MATLAB.

3. Results

Table 1 shows the classification rates achieved by each classification method for both subjects. Most methods tested are statistically better than the baseline (AR feature extraction /simple linear classification. The method with the highest mean classification rate is db25 filtering with polynomial SVM classification, which has an accuracy nearly 10% higher than the baseline. Figure 2 shows that a large improvement was achieved by

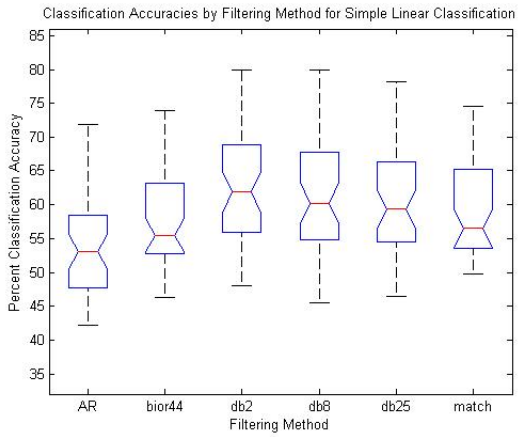

Figure 2. ***Classification rates for each filtering method and simple linear classification.***

using wavelets and the simple linear classifier.

4. Discussion

These results suggest that support vector machines with linear and nonlinear kernels are an effective way to classify EEG features. Nonlinear SVMs generally show a small advantage over linear SVMs, but the difference is slight. SVMs with all kernels statistically outperform the simple linear classifier. Further, they suggest that wavelets are indeed a good method for extracting movement-related information from EEG signals. Daubechies wavelets appear to be generally superior to biorthogonal wavelets, autoregressive filtering, and mu-matched filtering for BCI applications

Although the overall classification rates are low for all methods, it must be remembered that BCI applications process many samples per second. A typical BCI task may involve the processing of hundreds or even thousands of samples, and therefore it is only necessary that most samples are correctly interpreted for the BCI task to be successfully completed. Relatively small improvements in classification may therefore have large improvements in online BCI performance.

An important consideration for BCI applications is the need for processing algorithms to run in real time. Using an Intel Core Duo computer running at 2.66 GHz with 3.5 GB of RAM, AR, mu-matched, and db2 filtering were performed in real time with a negligible CPU cost. Db8 filtering was performed in real time, but consumed approximately 10% of the CPU resources, and db25 filtering was unable to be performed in real time. Therefore, db25 filtering is unlikely to be useful for online BCI processing without the use of a distributed algorithm using many processors, and db8 filtering may be problematic if limited computing resources are available.

SVM classification is very fast, but training a machine is slow, and for our application takes on the order of tens of minutes. If this is infeasible, dramatic gains can nevertheless be achieved by use of wavelets in combination with a simple linear classifier, as shown by figure 2.

References

[1] G. Fabiani, D. McFarland, J. Wolpaw, and G. Pfurtscheller, "Conversion of EEG activity into cursor movement by a brain-computer interface (BCI)," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on [see also IEEE Trans. on Rehabilitation Engineering]*, vol. 12, no. 3, pp. 331–338, 2004.

[2] E. Leuthardt, G. Schalk, J. Wolpaw, J. Ojemann, and D. Moran, "A brain–computer interface using electrocorticographic signals in humans," *Journal of Neural Engineering*, vol. 1, no. 2, pp. 63–71, 2004.

[3] E. Niedermeyer, L. da Silva, *et al.*, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins, 1999.

[4] D. Krusienski, G. Schalk, D. McFarland, and J. Wolpaw, "Tracking of the Mu Rhythm using an Empirically Derived Matched Filter," *Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference on*, pp. 86–89, 2005.

[5] R. Quian Quiroga, O. Sakowitz, E. Basar, and M. Schürmann, "Wavelet Transform in the analysis of the frequency composition of evoked potentials," *Brain Research Protocols*, vol. 8, no. 1, pp. 16–24, 2001.

[6] G. Schalk, D. McFarland, T. Hinterberger, N. Birbaumer, and J. Wolpaw, "BCI2000: a general-purpose brain-computer interface (BCI) system," *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 6, pp. 1034–1043, 2004.

[7] D. McFarland, L. McCane, S. David, and J. Wolpaw, "Spatial filter selection for EEG-based communication," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 3, pp. 386–394, 1997.

[8] D. Krusienski, D. Mcfarland, and J. Wolpaw, "An Evaluation of Autoregressive Spectral Estimation Model Order for Brain-Computer Interface Applications," *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, pp. 1323–1326, 2006.

[9] I. Daubechies, *Ten Lectures on Wavelets*. Society for Industrial Mathematics, 1992.

[10] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2000.

[11] C. Chang and C. Lin, "LIBSVM: a library for support vector machines," *Software available at http://www.csie. ntu. edu. tw/cjlin/libsvm*, vol. 80, pp. 604–611, 2001.

[12] C. Hsu, C. Chang, C. Lin, *et al.*, "A practical guide to support vector classification," *National Taiwan University, Tech. Rep., July*, 2003.